DOCUMENT RESUME

ED 155 212 TM 007 215

AUTHOR Crowder, Christopher B.; Gallas, Edwin J.
TITLE Relation of Out-of-Level Testing to Ceiling and Floor Effects on Third and Fifth Grade Students.
PUB DATE Mar 78
NOTE 43p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978); For related documents, see TM 007 212 and 216

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Achievement Tests; *Comparative Testing; Complexity Level; Elementary Education; Equated Scores; Grade 3; Grade 5; *Grade Equivalent Scores; Norm Referenced Tests; *Norms; Raw Scores; Reading Achievement; Reading Tests; *Scores; Standardized Tests; Student Reaction; *Testing; *Test Interpretation; Test Reliability
IDENTIFIERS Metropolitan Achievement Tests; *Out of Level Testing; Stanford Achievement Tests

ABSTRACT

Both on-level and out-of-level tests were administered to third and fifth grade children in order to compare the scaled scores of different level tests of the same testing program and to discover whether the relationship between levels might be distorted by ceiling or floor effects. Only reading tests were used in this study. The Stanford Achievement Tests were used for the third grade, and the Metropolitan achievement Tests were used for the fifth grade. The results are presented in tables and graphs, and the students' attitudes regarding difficulty of the tests are discussed. Overall, the comparative standard scale scores across the levels examined did not appear to interlock perfectly. Whether such measurement error was offset by reduction in the floor and ceiling effects for the out-of-level tests apparently depended heavily on two factors: the relative difficulty of the adjacent out-of-level test, and how much scores were initially depressed or inflated for the given group tested on-level. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED155212

Session No. 31.12

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Christopher R. Crowder

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM.

Presented at the Annual Meeting of the American Educational Research Association, Toronto, Canada, March 27-31, 1978

Relation of Out-of-Level Testing to Ceiling and Floor Effects on Third and Fifth Grade Students

Educators have criticized standardized tests in the past because grade level tests did not seem appropriate in content for students who previously scored at the extremes on such tests. More specifically, the wider the grade band covered by a test, the less likely it would be that the difficulty and content of the test would be suitable for students at the lower and upper ends of performance.

It has been suggested that out-of-level testing should occur when approximately three-fourths of the group tested can answer the most difficult items correctly. This is referred to as a ceiling effect. Conversely, when the mean is approximately one-third of a standard deviation above the median for the group tested, it is likely that floor effects have been encountered on the test. The avoidance of ceiling and floor effects produces a test that is said to be at a suitable or functional level for the students. The level used is considered to be appropriate when the mean raw score is equal to or above one-third of the maximum and less than three-fourths of the maximum score (Roberts, 1976).

Some Title I programs, such as in Rhode Island, have responded to such criticisms and suggestions by allowing local education agencies to test out-of-level rather than at grade level (on-level). The inherent assumption here is that scores derived from both levels for the same students would be comparable.

Previous studies have questioned this assumption, and one such study showed that generally, the higher the proportion of students tested out-of-level yearly, the lower the mean grade equivalent (Ayer and McNamara, 1973). Also, an investigation of out-of-level and on-level scores for the same children in the second, third and fourth grades demonstrated significant differences in grade equivalent scores for the two levels examined (Long, Schaffran and Kellogg, 1977). The trend for such grade equivalent score differences was not consistent across the three grades. However, the test used did not offer expanded standard scores, and thereby make it possible to relate the norms for the different test levels. Another study, tangentially related to the issue, determined that for equal raw scores achieved by out-of-level students at grades four, six, seven and eight in reading and mathematics, the item response patterns were comparable for the on-level, fifth grade students (Plake and Hoover, 1977).

This study investigated the following questions:

1. Are standard scale scores comparable for students between on-level and out-of-level tests, when the on-level test has previously shown ceiling effects?

2. Are standard scale scores comparable for students between on-level and out-of-level tests, when the on-level test has previously revealed floor effects?

3. Do students whose performance is in the guessing range for the on-level test also achieve at the guessing level on the out-of-level test?

4. Do students whose performance is in the ceiling range for the on-level test also achieve in the ceiling range on the out-of-level test?

5. Do students from schools which previously showed floor effects for the on-level test, find the out-of-level test easier and less frustrating?

6. Do students from schools which previously demonstrated ceiling effects for the on-level test, indicate that the out-of-level test is easier and less frustrating.

Method

Data for the study were collected from six schools for each grade in a metropolitan area within a single school district. Three of the schools were selected because their third grade students had previously exhibited ceiling effects on test results for the state mandated reading test, the Stanford Achievement Test, 1973 Edition, Primary Level II, Form A, normed for grades 2.0 through 3.9. Three schools were chosen because their fifth grade students had, in the past, demonstrated floor effects on test results for the district reading test, the Metropolitan Achievement Test, 1971 Edition, Intermediate Level, Form F, normed for grades 5.0 through 6.9. Within the six schools selected, eight classes were randomly chosen for the third grade study, with at least one class from each school. For the fifth grade study, one class was randomly selected from each of the six schools.

For each grade, third and fifth, the classes were randomly divided into two groups, one which took the regular on-level test

first, and one which took the out-of-level test first, to counterbalance the practice effects of the testing order. The out-of-level test administered in schools showing previous floor effects was one level lower than the on-level test, specifically the Primary Level I (normed for grades 1.5 through 2.9) for the third grade, and the Elementary level (normed for grades 3.0 through 4.9) for the fifth grade. Conversely, the out-of-level test administered in the schools showing previous ceiling effects was one level higher than the regular on-level test, namely the Primary Level III (normed for grades 3.9 through 4.9) for the third grade, and the Advanced level (normed for grades 6.9 through 9.0) for the fifth grade.

Each class was administered the on-level test at the regular time prescribed by the state or the district. Either one week before or one week afterwards, depending upon the random assignment, the class took the out-of-level test. Following the administration of both tests, the children were given a one-page questionnaire (see Appendix) concerning how they felt about the two tests. The questionnaire was read to the students by each teacher.

Results and Interpretation

For the third grade students, no significant ($p < .05$) practice order effects for testing were shown, either for groups showing previous floor effects or ceiling effects for the regular on-level test. Similarly, for fifth grade students, no significant ($p < .05$) order effects for the administration of the two test levels were demonstrated. Consequently, data for students tested initially or

last with the out-of-level test were combined for both the third grade and the fifth grade.

The estimated guessing level established for the study was one-fourth of the points possible (four response items) which is conservative, since one or two standard deviations above this estimate could be interpreted as signifying no knowledge of the subject matter (Gulliksen, 1950). Also, the estimated ceiling level was established at three-fourths of the maximum points (Roberts, 1976). Scores falling between one-third of the maximum score and the ceiling level were considered to be in-range, or suitable, for the test.

Table 1 for third grade students, based upon raw scores, shows several points of interest. First, in reference to the suitability of each test level, it can be observed that the mean raw score is above one-third and less than three-fourths of the maximum raw score for students from schools previously showing floor effects on both the on-level (Primary II) and out-of-level (Primary I) tests. Second, for the students from schools previously showing ceiling effects, the on-level raw score mean is above the suitable or functional range, but the out-of-level mean raw score is within such a range. Also, the on-level mean is more than one-third of a standard deviation below the median which tends to confirm that the on-level test is not suitable for these students (Roberts, 1976).

The most reliable mean score appears for the out-of-level Primary I test, where the students answered somewhat more than one-half of the items correctly (Roberts, 1976).

Insert Table 1 about here

For fifth grade students, Table 2 reveals a different pattern. The mean scores for students tested out-of-level (Elementary and Advanced) fall within a functional or suitable range, as does the raw score mean on the on-level test (Intermediate) for students from schools showing previous ceiling effects. However, the on-level mean score for students from schools showing previous floor effects falls below the suitable range. Also, the most reliable mean score is found for students from previous ceiling effect schools for the on-level test (Intermediate). The median score for this group is also well above half the number of test items for the on-level test, but it is not above half for either the out-of-level test nor the on-level test for the previously floor effect schools.

Insert Table 2 about here

In summary, Table 1 supports the contention that the on-level reading test is relatively easy for third grade students from schools previously showing ceiling effects. Table 2 upholds the hypothesis that the on-level reading test is relatively hard for fifth grade students from schools previously indicating floor effects. Both conditions set the stage for comparing standard scale scores between on-level tests (which demonstrate either floor or ceiling effects) and out-of-level tests.

A comparison in Table 3 of the standard scale scores for the third grade between the on-level and out-of-level tests shows no significant differences ($p < .05$) overall, for either students tested with the lower out-of-level test (Primary I) or students tested with the higher out-of-level test (Primary III). The mean standard scale score differences were less than two units for both groups.

Insert Table 3 about here

Table 5 shows the correlations and the coefficients of determination between each test level on the third grade Stanford Reading Test. It reveals that approximately 60 percent of the variance of the Primary I level is associated with the variability of the Primary II level. A similar relationship holds for the Primary II and Primary III levels. The fairly large amount of variance common to the on-level and out-of-level tests supports the small mean scale score differences found between the two levels, as indicated in Table 3.

Insert Table 5 about here

For the fifth grade Metropolitan Reading Test, a comparison between on-level and out-of-level standard scale scores in Table 4 reveals significant differences ($p < .05$). For both the students at schools previously showing floor effects and those at schools

formerly demonstrating ceiling effects, the mean standard scores favored the on-level results.

Insert Table 4 about here

The variance common to the on-level and out-of-level tests for the fifth grade Metropolitan Reading Test is considerably less than the common variance accounted for between levels on the third grade reading test. Table 6 reveals that from 32 to 46 percent of the variance is associated between the fifth grade on-level and out-of-level tests. This parallels the significant differences between levels for the mean standard scores as disclosed in Table 4.

Insert Table 6 about here

In general, for both the third grade students and the fifth grade students, on the Stanford and Metropolitan Reading Tests respectively, the standard scale score results do not favor testing out-of-level. This holds for schools previously showing floor effects for the on-level test and therefore testing one level below the on-level test. It also is true for schools formerly revealing ceiling effects for the on-level test and consequently testing one level above the on-level test.

More specifically, how do students perform when assessed on the next lower level of the test, if they scored at the guessing level or below (show floor effects) for the on-level test?

Similarly, how do students score on the next higher level of the test, if they achieved at or above the ceiling for the on-level test?

Tables 7 through 10 show the mean standard scale scores for students who scored at the guessing level or lower for the on-level test.

Table 7 specifically indicates that third grade students who scored at the guessing level or lower for the on-level test (Primary II) achieved higher standard scale scores on the next lower level of the test (Primary I). The in-range scores were comparable for the two levels, and students achieving at or above ceiling for the on-level test scored slightly higher on the lower test level. It should be noted that these data tend to confirm that the state mandated Stanford Reading Achievement Test, Primary Level II, Form A, is relatively easy. Even students selected from schools achieving in the lower 15 percent of the district on this test, tend to score at a functional level or above. Nevertheless, the next lower level of the test offers an advantage to students scoring at the guessing level or lower on the on-level test.

Insert Table 7 about here

Table 8 discloses a somewhat different pattern for third grade students selected from schools performing in the upper 15 percent of the district. No significant gain is observed for students tested at the next higher level (Primary III) who scored in-range for the

on-level test. Similarly, no significant increase occurs at the next higher level for students who scored in the ceiling range for the on-level test.

Insert Table 8 about here

Data for the district administered fifth grade test, the Metropolitan Reading Test, show that it tends to be a more difficult measure than the third grade reading test. Table 9 reveals that for students from schools performing in the lower 15 percent of the district for the on-level test (Intermediate level), no students achieved standard scale scores at the ceiling level or above. Interestingly enough, the same students did not score in the ceiling range on the next lower level of the test (Elementary level). For such fifth grade students performing in the functional range for the on-level test, a significant average decrease in standard scale scores occurs on the next lower level. Students scoring in the guessing range for the on-level test still averaged higher standard scores than for the next lower level of the test. No advantage occurs, in reference to standard scale scores, in testing these students with the next lower level.

Insert Table 9 about here

For students from schools that achieved in the upper 15 percent of the district for the on-level test, Table 10 shows that fifth grade students scoring in-range averaged significantly lower scores on

the next higher level of test (Advanced level). This pattern also held for students scoring in the ceiling range for the on-level test. The standard score differences favor the use of the on-level test for such students, except for the two fifth grade subjects who scored in the guessing range for the on-level test.

Insert Table 10 about here

Generally, for both third and fifth grade students, no significant gains in standard scale scores occur for students at different functional levels of the test (floor, in-range and ceiling), when they are tested with the next lower or higher level of the test. The one significant exception to this pattern develops for third grade students scoring in the guessing range (N=14) for the on-level Stanford Reading Test. For the apparently more difficult fifth grade Metropolitan Reading Test, no real advantage in standard scale scores develops by testing out-of-level. Of more importance perhaps, no out-of-level gains in standard scale scores result for the 26 students scoring in the guessing range on the relatively difficult fifth grade on-level test.

Another way to look at the differences in on-level and out-of-level test results is by comparing the number and percentage of students scoring at the guessing, in-range and ceiling levels for each test level.

Table 11 indicates that for all third grade students tested, either at the next lowest or the next highest level, a greater

percentage scored in the guessing floor and ceiling range on the on-level test (Primary II) than on the out-of-level test (Primary I). Conversely, a larger percentage scored at the functional level (in-range) on the out-of-level test.

Insert Table 11 about here

For the fifth grade reading test, as revealed in Table 12, there is very little difference in the percentage of students scoring in the guessing, functional and ceiling ranges for the on-level and out-of-level tests. However, the pattern follows the third grade differences between levels, since slightly more fifth grade students score at the floor and ceiling range for the out-of-level test (Advanced) than the on-level test (Intermediate).

Insert Table 12 about here

What appears significant, when comparing Tables 7 through 10 to Tables 11 and 12, is that differences in the mean performance for the students scoring in the guessing, functional (in-range) and ceiling range do not parallel differences in the percentage of students scoring in each of these performance ranges for the on-level and out-of-level tests. This is apparent on test results for both the third and fifth grade students, and the question is--why?

If the raw scores for each level of test are plotted against the standard scale scores, as shown in Figures 1 through 4 for the third grade students, it can be observed that the distribution of

scores is not linear at the upper score extremes (in the ceiling range) for Figures 1 and 3. Also, relatively large increases in standard scale scores occur for small raw score gains at such extremes. In contrast, in the functional range, several raw scores have the same standard scale score.

Insert Figures 1-4 about here

Table 13 supports this finding, as the Pearson product-moment correlation coefficients are less than 1.00 between the raw score and standard score for each school group at each test level, and especially for the levels where ceiling effects are most frequent. In comparison, the Spearman rank-difference coefficients approach 1.00 for each test level and both groups, since this statistic does not take into account the quantitative deviation of each score from the mean (Guilford and Fruchter, 1973).

Insert Table 13 about here

Raw scores plotted against standard scale scores for fifth grade students on each level of the Metropolitan Reading Test, as shown in Figures 5 through 8, reveal a rather different pattern of score distributions than for the third grade data. Here, the distributions of scores approach linearity except at the extremes of the guessing and ceiling range, as shown in Figures 6 and 7.

Insert Figures 5-8 about here

Although relatively small increases in raw scores produce large gains in standard scale scores at such extremes, few raw scores offer the same standard scale score. Table 14 discloses a much closer agreement between the Pearson product-moment and Spearman rank-difference coefficients, and also reflects the raw score--standard scale score distributions for the levels of this test.

Insert Table 14 about here

Table 15 demonstrates the linear and curvilinear regression analyses of raw scores on standard scale scores for the on-level and out-of-level third grade reading tests. It can be seen that the variance explained by the curvilinear (quadratic) regression is relatively greater for the out-of-level (Primary I) test and the on-level (Primary II) test for students from schools previously showing ceiling effects. Both Figures 1 and 3 support the ceiling effects for these two tests.

Insert Table 15 about here

Linear and curvilinear analyses for the fifth grade reading tests are shown in Table 16. Most of the variance is explained by linear regression, and only for the on-level (Intermediate) test

for students from schools formerly demonstrating floor effects does the curvilinear analysis explain much more of the score variance. Figure 6 tends to confirm the floor effect for these students.

Insert Table 16 about here

Do out-of-level tests make the student from schools which tend to show floor or ceiling effects for the on-level test feel better about their performance? Perhaps an attempt to directly assess such feeling is not as easy as it may appear initially. A questionnaire was used in this study to obtain information on the students' relative perceptions of both the on-level and out-of-level test. The Appendix shows the questions asked of each student. The instrument was read by the teacher to the students in each participating class (although the form says "may read"), but the questions appeared to be difficult for many of the third grade children to comprehend.

Questions one and two were so constructed that the student should not respond "yes" to both questions or "no" to both questions. Also the responses to questions three and four should parallel the student's response to question one. Responses to questionnaires which did not follow such patterns were considered unreliable and invalid. Questions five and six, although not intentionally so, appear to be ambiguous. These questions about guessing may reflect more than one underlying behavior or attitude. The student may respond to these questions in reference to the relative difficulty of the test or

rather in terms of their test-taking strategy (guessing). For this reason, the responses to questions five and six have not been included in the results. Questions seven and eight, not directly related to test level attitude, were included to determine the student's general attitude toward reading.

Considering only questionnaires with reliable response patterns, the following third and fifth grade results were obtained. For third grade students taking the next lower level of test for out-of-level (N=23), 100 percent believed that the out-of-level test was easier than the on-level test. Also all of these students indicated that they knew more words and more answers in the out-of-level test. Ninety-six percent stated they liked to read to learn things, and 78 percent responded that they liked to read when they have extra time.

Third grade students who took the next higher out-of-level test produced some interesting responses. Although one teacher stated that the on-level test was easier for the students, the reliable student responses do not confirm this. Eighty-eight percent (N=17) marked that the next higher, out-of-level test was easier than the on-level test. The same percentage of students indicated that they knew more words and answers to the out-of-level test. All of these students said they liked to read to learn things, and 88 percent stated that they read when they have extra time.

The general favoring of the out-of-level test by the third grade students, over the on-level test appears analogous to the relative percentage of students performing in the functional range for the out-of-level test, as shown in Table 9. The response patterns support the data in Table 1, indicating that the higher out-of-level test is more appropriate.

Fifth grade students who were administered the next lower out-of-level test (N=32), were evenly divided on whether the on-level or out-of-level test was easiest. Fifty percent of the students also indicated that they knew more words and answers to each level of the test. All of them said they liked to read to learn things, and 75 percent revealed they enjoyed reading when they have extra time.

Ninety-five percent of the fifth grade students who took the next higher out-of-level test (N=37) reported that the on-level test was easier than the out-of-level test. The same percentage also marked that they knew more words and answers for the on-level test. Eighty-nine percent revealed that they like to read to learn things, and 87 percent said they like to read when they have extra time.

The mixed responses for students who took the next lower level may reflect the data in Table 7, which reveal that no students scored in the ceiling range for either the on-level or next lower level of the test. Also responses for students taking the next higher level parallel the relative performance of such students as

shown in Table 8. Similarly, the questionnaire results support the data in Table 2 indicating that the on-level test is relatively difficult.

Conclusions

The answer to how comparable the standard scale scores are between the on-level and out-of-level tests appears related more to the test being used, its raw score-standard score scale transformation and the relative difficulty of the test, than to the previous floor or ceiling effects shown by students tested on-level. Results for third grade students on the Stanford Reading Test show no real differences between the on-level test and the next higher or next lower level. However, a study of mean raw scores for the different test levels indicates that, for students from schools previously showing ceiling effects, the on-level test mean is in the ceiling rather than the functional or suitable range. The amount of common variance between the levels is relatively high.

In contrast, the fifth grade Metropolitan Reading Test reveals significant differences for students between the on-level and out-of-level tests both from schools previously showing floor effects and ceiling effects. The standard score differences, in both instances, favor testing on-level. However, for the on-level test, the raw score mean is below the suitable range for students from schools previously showing floor effects. The relative lack of common variance between the levels supports these findings.

The answer to the more specific question whether students who perform in the guessing and ceiling ranges for the on-level test

also achieve in the same ranges on the out-of-level tests appears more complex. For the third grade reading test, students from schools previously showing floor effects and individually scoring in the guessing range for the on-level test, did achieve significantly higher standard scores on the next lower test level. However, no real difference was observed between test levels for students from schools previously showing ceiling effects, who individually achieved in the ceiling range for the on-level test.

For students from schools previously showing floor effects for the on-level fifth grade reading test, individuals scoring in the guessing range on-level did not perform better on the next lower level. Fifth grade students from schools formerly indicating ceiling effects for the on-level test, who individually scored in the ceiling range on-level, did not achieve higher standard scores out-of-level.

Generally, both in third and fifth grade, students did not achieve significantly higher standard scores in reading on the out-of-level tests. A study of the percentage of students scoring in the guessing and ceiling ranges for the on-level and out-of-level tests, however, produced different results and possible conclusions.

Compared to the on-level third grade reading test, fewer students scored in the guessing and ceiling ranges on the out-of-level tests. For the fifth grade reading test, no real differences in percentage of students scoring in the guessing and ceiling

ranges were observed between the levels of the test. Why are these percertage results not comparable to the mean scale scores between test levels?

The differences in mean scale scores and percentages for the on-level and out-of-level tests may be partially accounted for by the relatively large increase in standard scores for raw scores at the extremes of the scale transformation. This would occur when scores in the guessing or ceiling ranges appear frequently. Also, on one scale especially--for the Stanford Reading on-level test, several raw scores convert to the same standard score. Correlations between the two types of scores support both observations. Tests of linearity and curvilinearity support the presence of ceiling effects for levels of the third grade reading test, and some floor effects for the fifth grade reading test.

Students taking the out-of-level tests for third grade reading appeared to find such tests easier, even those students taking the next higher level. These responses support the larger percentage of students scoring in the suitable or functional level on the out-of-level tests.

Fifth grade students taking the next lower level of test were evenly divided on which test level was easier, but students taking the next higher out-of-level test reported that the on-level test was easier. These responses tend to confirm the relative difficulty of the on-level test, and the fact that no additional students scored in the ceiling range on the next lower test level.

Overall, the comparative standard scale scores across the levels examined do not appear to interlock perfectly. Whether such measurement error is offset by reduction in the floor and ceiling effects for the out-of-level tests apparently depends heavily on two factors: (a) the relative difficulty of the adjacent out-of-level test and (b) how much scores are initially depressed or inflated for the given group tested on-level. The findings are restricted to the tests used and the students at the grades studied.

References

Ayer, J.E., & McNamara, T.C. Survey testing on a out-of-level basis. Jornal of Educational Measurement, summer 1973, 10(2), 79-83.

Guilford, J.P. & Fruchter, B. Fundamental statistics in psychology and education. New York: McGraw-Hill, 1973.

Gulliksen, H. Theory of mental tests. New York: John Wiley & Sons, Inc., 1950.

Long, J.V., Schaffran, J.A., & Kellogg, T.M. Effects of out-of-level survey testing on reading scores of Title I, ESEA students. Journal of Educational Measurement, fall 1977, 14(3), pp. 208-213.

Plake, B.S., & Hoover, H.D. The comparability of equal raw scores for children at different grade levels: an issue in "out-of-level" testing. Paper presented at NCME annual meeting, New York, April 1977.

Roberts, A.O.H. Out-of-level testing. ESEA Title I Evaluation and Reporting System (Technical Paper No. 6). Mountain View, California: RMC Research Corporation, October 1976.

Table 1

Comparison of Raw Score Means, Medians and Standard Deviations for On-Level and Out-of-Level Tests of Stanford Reading on Third Grade Students

Test Level	School Group	N	Items	Mean	Median	Standard Deviation
Primary I	Floor	75	87	58.96	63.58	19.17
Primary II	Floor	75	93	42.61	37.12	17.49
Primary II	Ceiling	92	93	80.54	85.64	11.96
Primary III	Ceiling	92	70	49.15	49.50	12.70

Table 2

Comparison of Raw Score Means, Medians and Standard Deviations for On-Level and Out-of-Level Tests of Metropolitan Reading on Fifth Grade Students

Test Level	School Group	N	Items	Mean	Median	Standard Deviation
Elementary	Floor	80	45	18.90	17.10	8.20
Intermediate	Floor	80	45	14.18	13.50	5.39
Intermediate	Ceiling	49	45	26.59	27.00	8.47
Advanced	Ceiling	49	45	18.29	17.25	7.24

Table 3

Comparison of Standard Scale Scores for Third Grade Students Between On-Level and Out-of-Level Results on the Stanford Reading Test

School	Test Level	N	Mean	SE_M	Mean Difference	t[1]
Floor Group						
	Primary I	75	117.01	1.71	1.17	1.13
	Primary II	75	115.84	1.52		
Ceiling Group						
	Primary II	92	145.05	1.40	1.23	1.25
	Primary III	92	146.28	1.57		

[1]Note: Correlated t tests were used to compare mean differences for significance.

Table 4

Comparison of Standard Scale Scores for Fifth Grade Students Between On-Level and Out-of-Level Results on the Metropolitan Reading Test

School	Test Level	N	Mean	SE_M	Mean Difference	t[1]
Floor Group						
	Elementary	80	55.78	1.38	7.50	6.29**
	Intermediate	80	63.28	1.13		
Ceiling Group						
	Intermediate	49	80.88	1.67	3.33	2.42*
	Advanced	49	77.55	1.73		

[1]Note: Correlated t tests were used to compare mean differences for significance.

* $p < .05$
** $p < .01$

Table 5

Pearson Product-Moment Correlations and Coefficients of Determination for Standard Scale Scores Between Levels of the Third Grade Reading Test

Test Levels	N	r	r^2
Primary I vs. Primary II	75	.800	.640
Primary II vs. Primary III	92	.786	.618

Table 6

Pearson Product-Moment Correlations and Coefficients of Determination for Standard Scale Scores Between Levels of the Fifth Grade Reading Test

Test Levels	N	r	r^2
Elementary vs. Intermediate	80	.568	.322
Intermediate vs. Advanced	49	.675	.456

Table 7

Mean Standard Scale Scores on the Primary I and Primary II Levels by Performance Ranges for the On-Level Stanford Reading Test for Third Grade Students

Test Level	Floor N	Floor Mean	In-Range N	In-Range Mean	Ceiling N	Ceiling Mean	Total N	Total Mean
Primary I	14	103.07	52	116.59	9	141.11	75	117.01
Primary II	14	96.14	52	117.31	9	138.00	75	115.84
Mean Difference		6.93*		-.72		3.11		1.17

Correlated t tests were used to compare mean difference for significance

*$p < .05$

Table 8

Mean Standard Scale Scores on the Primary II and Primary III Levels by Performance Ranges for the On-Level Stanford Reading Test for Third Grade Students

Test Level	Floor N	Floor Mean	In-Range N	In-Range Mean	Ceiling N	Ceiling Mean	Total N	Total Mean
Primary II	0		14	126.42	78	148.40	92	145.05
Primary III	0		14	126.71	78	149.79	92	146.28
Mean Difference				-.29		-1.39		-1.23

Table 9

Mean Standard Scale Scores on the Elementary and Intermediate Levels by Performance Ranges for the On-Level Metropolitan Reading Test for Fifth Grade Students

Test Level	Floor N	Floor Mean	In-Range N	In-Range Mean	Ceiling N	Ceiling Mean	Total N	Total Mean
Elementary	26	48.39	54	59.33	0		80	55.78
Intermediate	26	51.65	54	68.87	0		80	63.28
Mean Difference		-3.26		-9.54**				-7.50**

**p < .01

Table 10

Mean Standard Scale Scores on the Intermediate and Advanced Levels by Performance Ranges for the On-Level Metropolitan Reading Test for Fifth Grade Students

Test Level	Floor N	Floor Mean	In-Range N	In-Range Mean	Ceiling N	Ceiling Mean	Total N	Total Mean
Intermediate	2	53.00	35	77.89	12	94.25	49	80.88
Advanced	2	69.50	35	73.54	12	90.58	49	77.55
Mean Difference		-16.50		4.35**		3.67		3.33*

*p < .05

**p < .01

Table 11

Comparison of Percentage of Students Testing in Three Ranges of Performance on the On-Level and Out-of-Level Tests for Third Grade Reading

Performance Range	On-Level N	On-Level %	Out-of-Level N	Out-of-Level %	Difference N	Difference %
Floor	14	8.4	4	2.4	10	6.0
In-Range	66	39.5	91	54.5	-25	-15.0
Ceiling	87	52.1	72	43.1	15	9.0
Total	167	100.0	167	100.0		

Table 12

Comparison of Percentage of Students Testing in Three Ranges of Performance on the On-Level and Out-of-Level Tests for Fifth Grade Reading

Performance Range	On-Level N	On-Level %	Out-of-Level N	Out-of-Level %	Difference N	Difference %
Floor	28	21.7	25	19.4	3	2.3
In-Range	89	69.0	95	73.6	-6	-4.6
Ceiling	12	9.3	9	7.0	3	2.3
Total	129	100.0	129	100.0		

Table 13

Comparison of Correlation Coefficients Calculated by Pearson Product-Moment and Spearman Rank-Difference Methods Between Raw Score and Standard Scale Score for Three Levels of Third Grade Reading Test

Test Level	School Group	N	r	p
Primary I	Floor	75	.902	.999
Primary II	Floor	75	.952	.999
Primary II	Ceiling	92	.850	.999
Primary III	Ceiling	92	.978	.999

Table 14

Comparison of Correlation Coefficients Calculated by Pearson Product-Moment and Spearman Rank-Difference Methods Between Raw Score and Standard Scale Score for Three levels of Fifth Grade Reading Test

Test Level	School Group	N	r	p
Elementary	Floor	80	.988	1.000
Intermediate	Floor	80	.972	1.000
Intermediate	Ceiling	49	.989	.999
Advanced	Ceiling	49	.973	1.000

Table 15

Results of Regression of Raw Scores on Standard Scores for Third Grade Reading Test

	df*	Sum of Squares	F	Significance of F	Multiple r	Variance Explained
Primary I Level Raw Scores for Floor Group						
Linear	1	22107.5	317.1	.0001	.90	.81
Quadratic	1	3699.8	191.7	.0001	.97	.14
Residual	72	1389.6	--	--	--	.05
Primary II Level Raw Scores for Floor Group						
Linear	1	20495.8	700.5	.0001	.95	.91
Quadratic	1	1748.2	324.5	.0001	.99	.08
Residual	72	387.8	--	--	--	.01
Primary II Level Raw Scores for Ceiling Group						
Linear	1	9407.4	235.1	.0001	.85	.72
Quadratic	1	3151.0	622.7	.0001	.98	.24
Residual	89	450.4	--	--	--	.04
Primary III Level Raw Scores for Ceiling Group						
Linear	1	14032.9	1964.1	.0001	.98	.95
Quadratic	1	539.4	463.1	.0001	.99	.04
Residual	89	103.7	--	--	--	.01

*degrees of freedom

Table 16

Results of Regression of Raw Scores on Standard Scores for Fifth Grade Reading Test

	df*	Sum of Squares	F	Significance of F	Multiple r	Variance Explained
Elementary Level Raw Scores for Floor Group						
Linear	1	5191.9	3283.6	.0001	.99	.98
Quadratic	1	81.7	151.1	.0001	.99	.01
Residual	77	41.6	--	--	--	.01
Intermediate Level Raw Scores for Floor Group						
Linear	1	2166.9	1313.4	.0001	.97	.94
Quadratic	1	108.9	422.4	.0001	.99	.05
Residual	77	19.8	--	--	--	.01
Intermediate Level Raw Scores for Ceiling Group						
Linear	1	3263.6	850.9	.0001	.97	.95
Quadratic	1	12.4	3.4	.072	.98	.01
Residual	46	167.9	--	--	--	.04
Advanced Level Raw Scores for Ceiling Group						
Linear	1	2467.6	2302.6	.0001	.99	.98
Quadratic	1	38.0	141.6	.0001	.99	.01
Residual	46	12.4	--	--	--	.01

*degrees of freedom

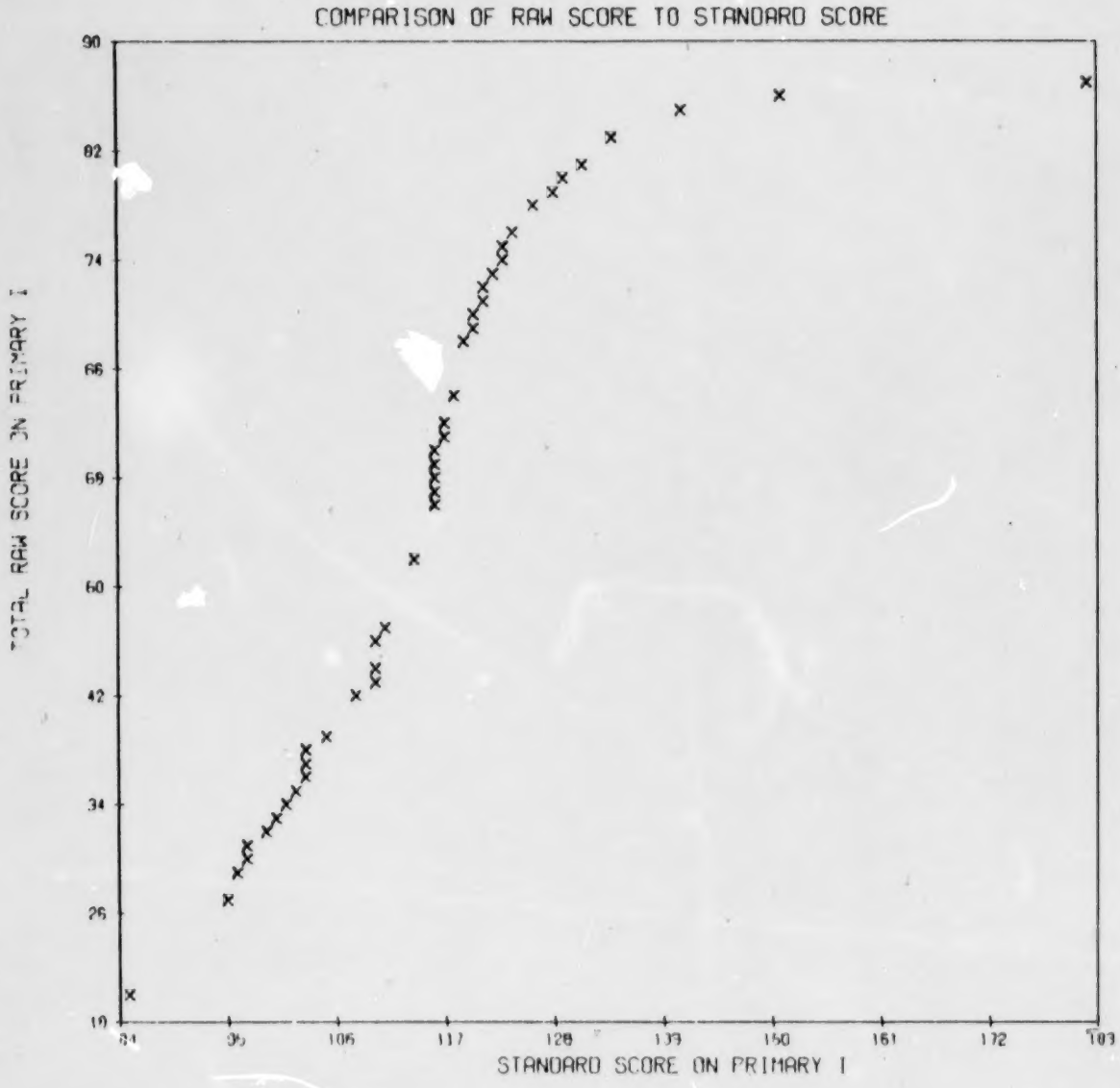

Figure 1. Comparison of Raw Score to Standard Score for the Primary I Level of the Third Grade Stanford Reading Test.

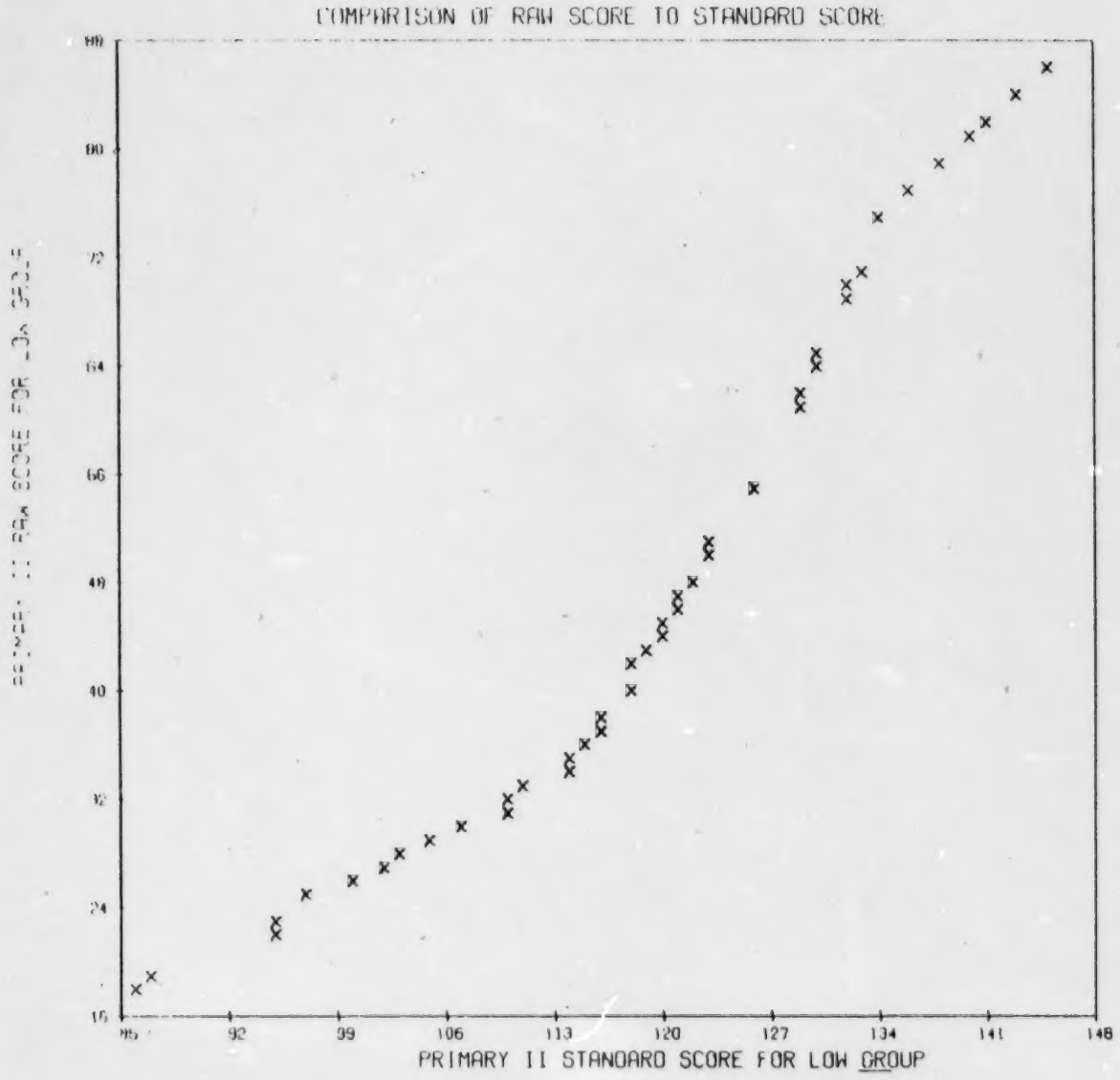

Figure 2. Comparison of Primary II Level Raw Score to Standard Score for Students from Schools Previously Showing Floor Effects on the Third Grade Reading Test.

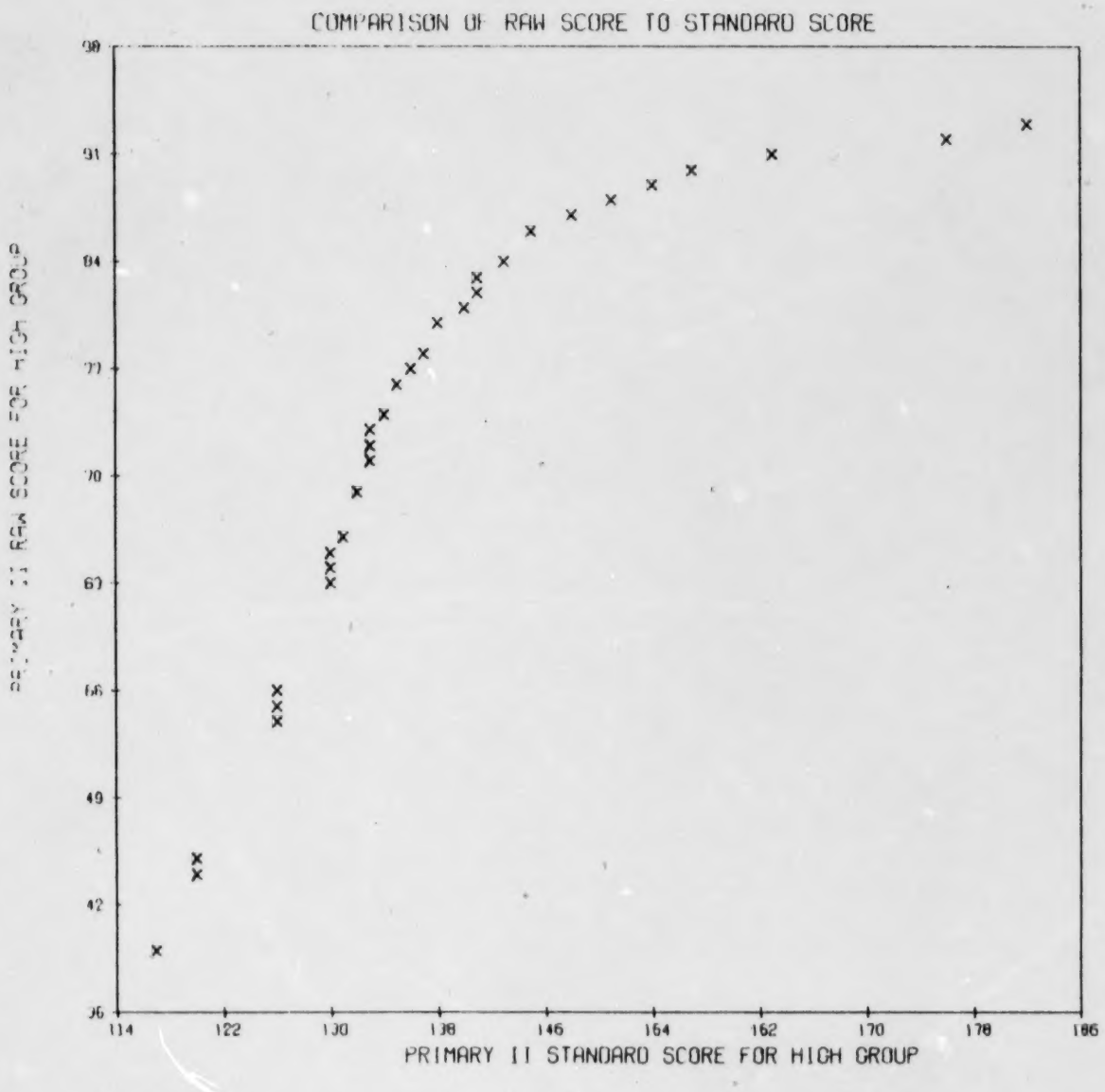

Figure 3. Comparison of Primary II Level Raw Score to Standard Score for Students from Schools Previously Showing Ceiling Effects on the Third Grade Reading Test.

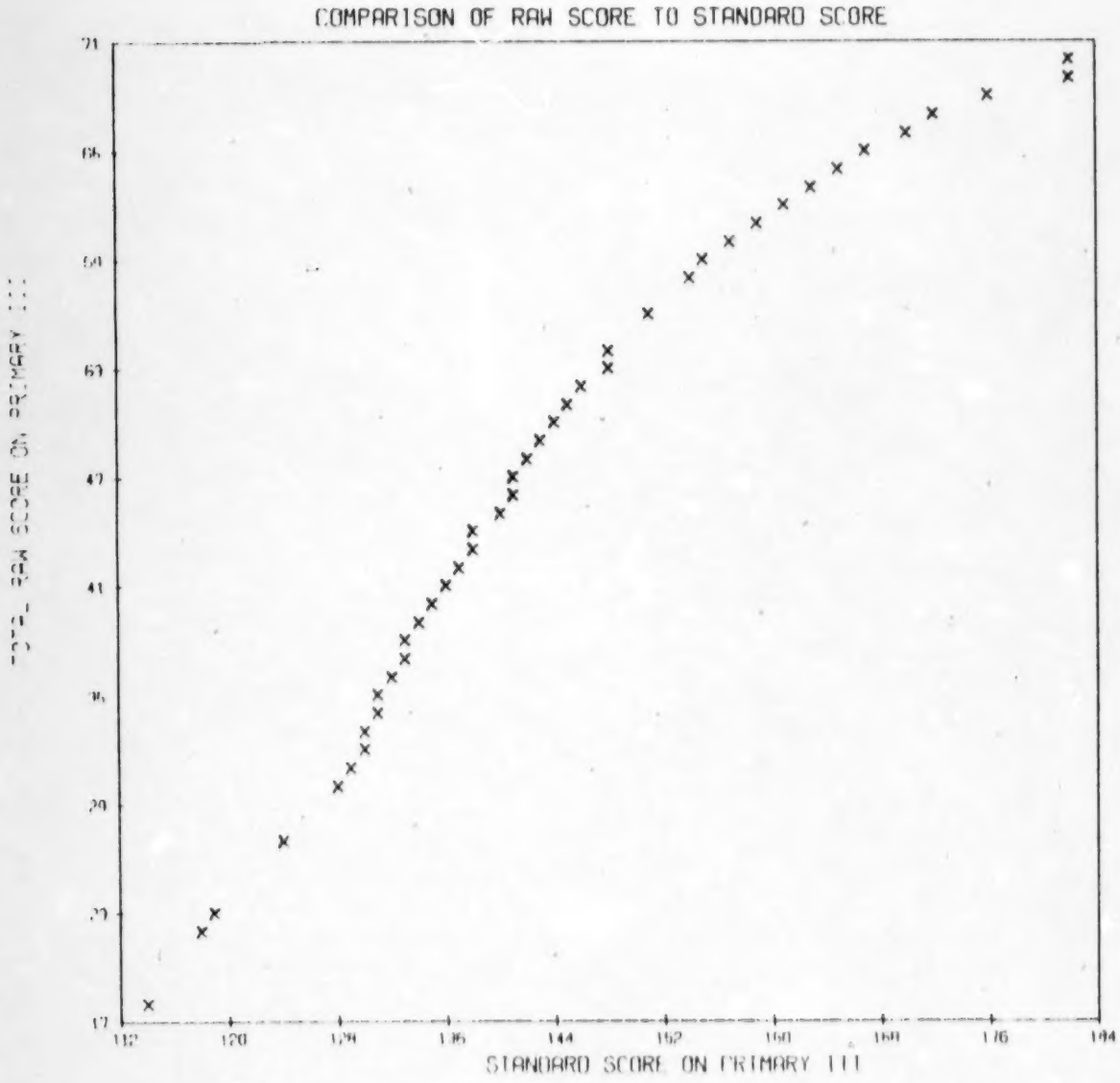

Figure 4. Comparison of Raw Score to Standard Score for the Primary III Level of the Third Grade Stanford Reading Test.

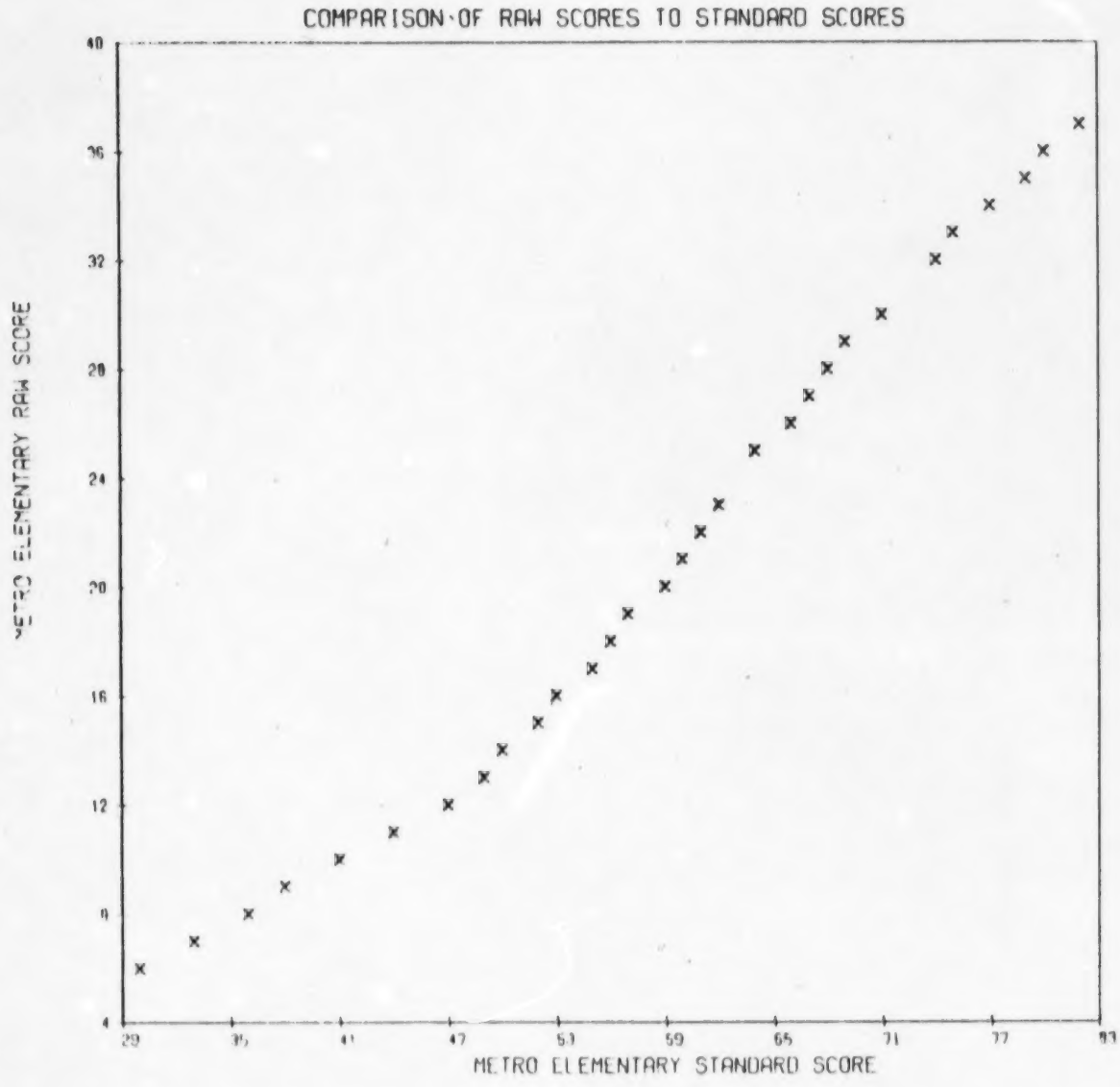

Figure 5. Comparison of Raw Score to Standard Score for the Elementary Level of the Fifth Grade Metropolitan Reading Test.

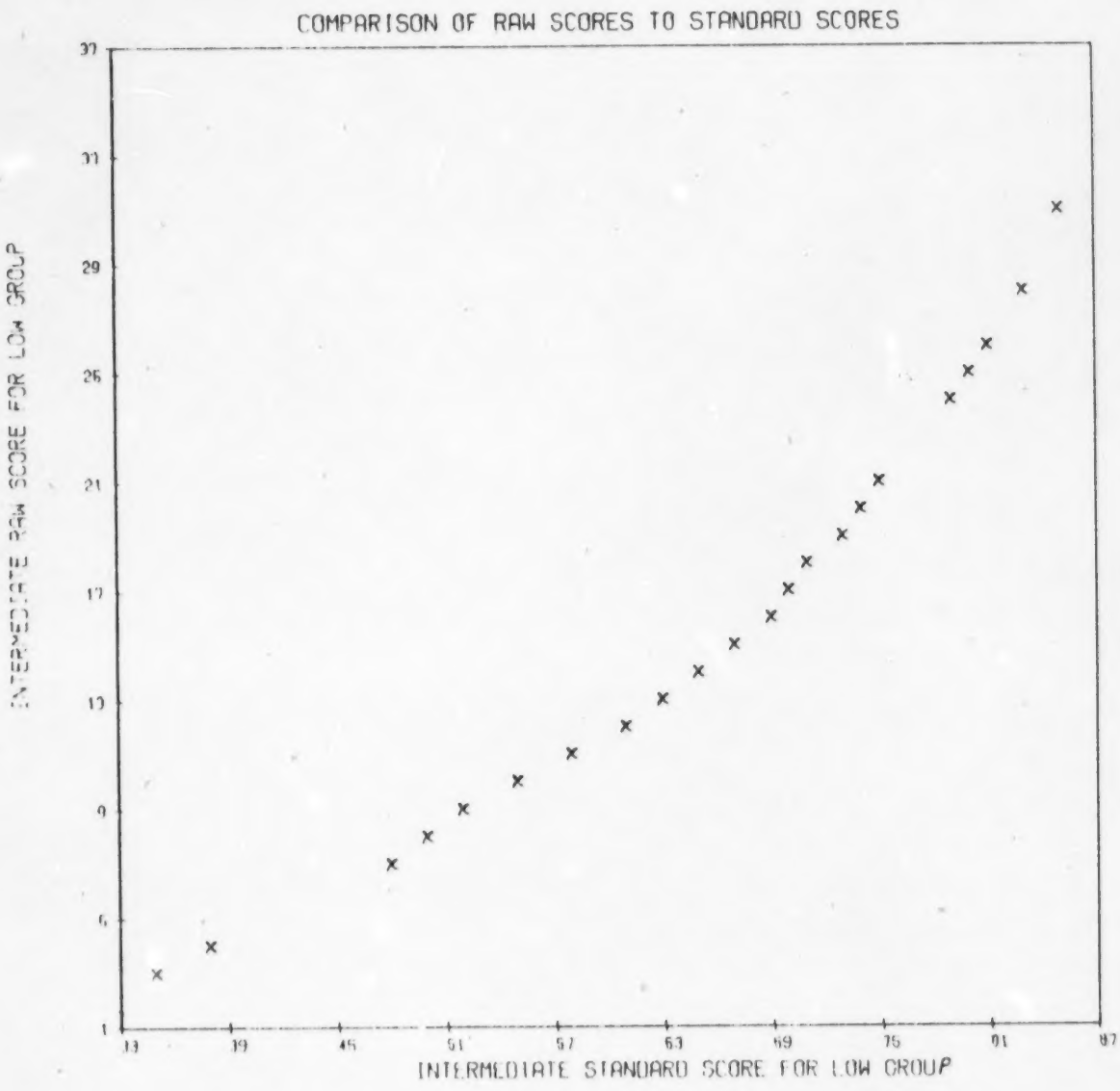

Figure 6. Comparison of Intermediate Level Raw Score to Standard Score for Students from Schools Previously Showing Floor Effects on the Fifth Grade Reading Test.

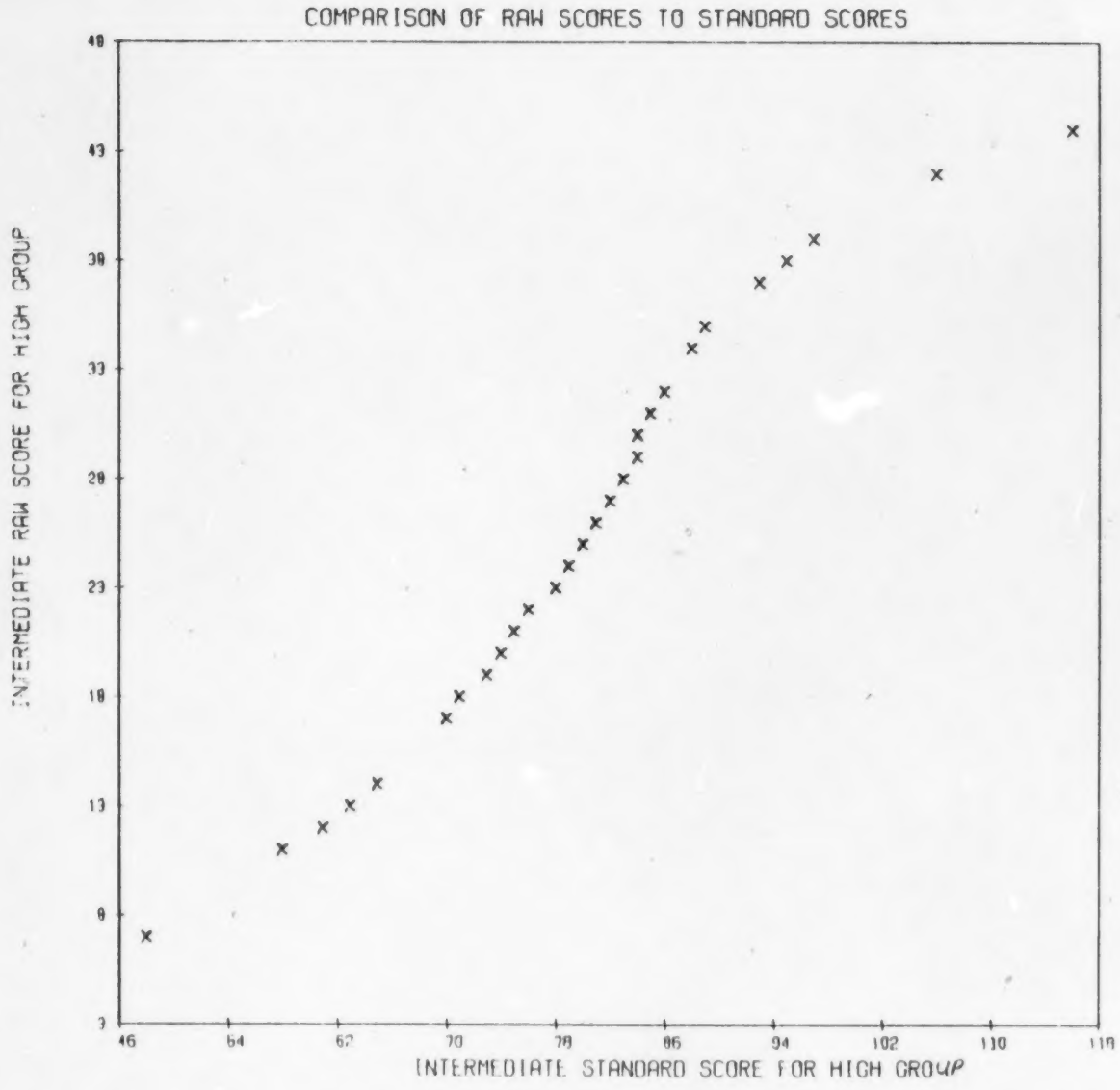

Figure 7. Comparison of Intermediate Level Raw Score to Standard Score for Students from Schools Previously Showing Ceiling Effects on the Fifth Grade Reading Test.

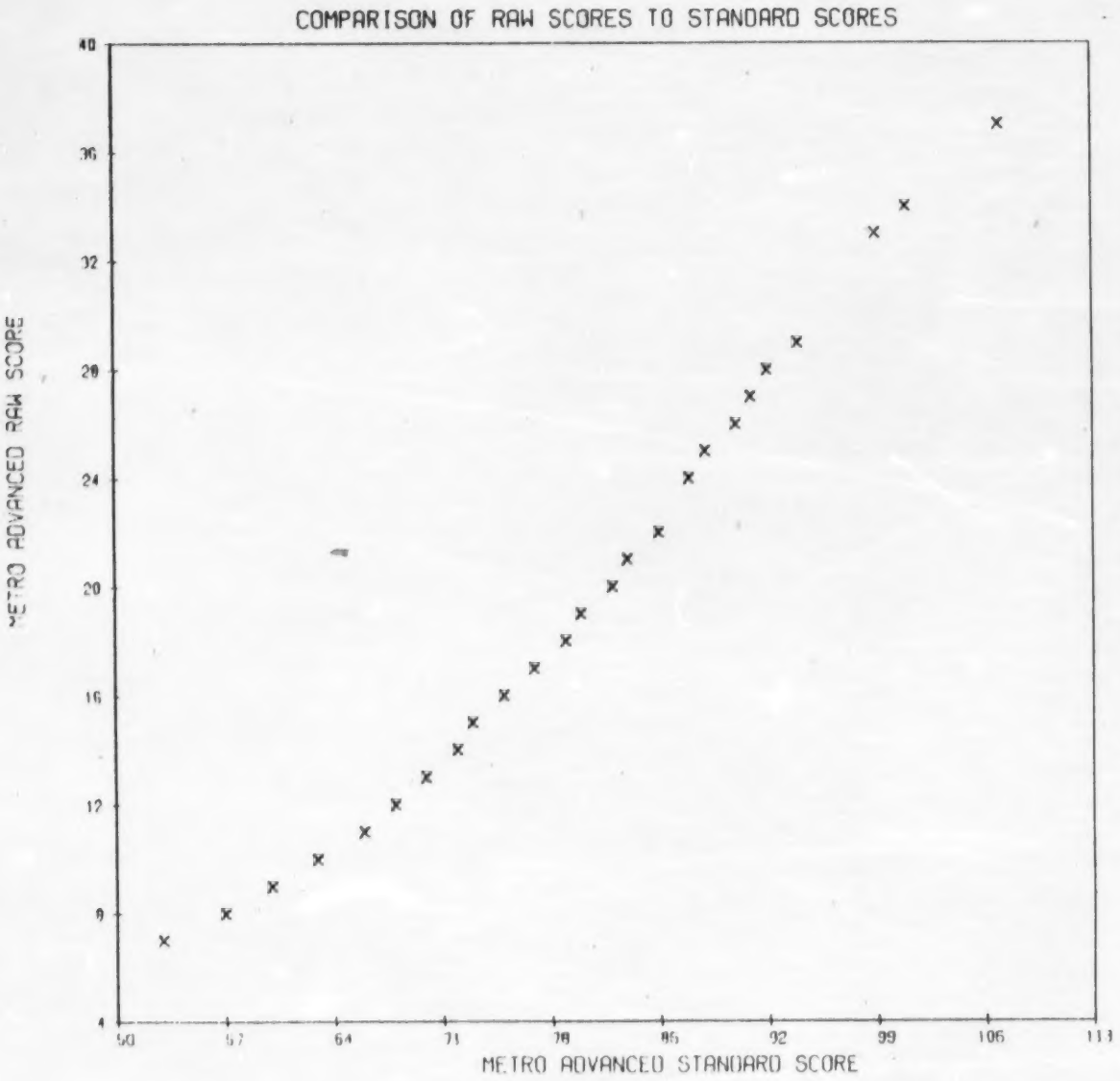

Figure 8. Comparison of Raw Score to Standard Score for the Advanced Level of the Fifth Grade Metropolitan Reading Test.

APPENDIX

Name_____ Matric_____ Birthdate_____

Grade_____ School_____ Teacher_____

HOW I FEEL ABOUT READING TESTS

Directions: The following questions are to find out how you feel about the two reading tests you have recently taken. Read* each statement and decide how you feel about it. Put an "X" in the space under "Yes" or "No" next to the question number, depending on how you feel.

Yes	No		
___	___	1.	The first reading test was easier than the second reading test.
___	___	2.	The second reading test was easier than the first reading test.
___	___	3.	I knew more of the words in the first reading test than in the second reading test.
___	___	4.	I knew more of the answers in the first reading test than in the second reading test.
___	___	5.	I guessed on many of the questions on the first reading test.
___	___	6.	I guessed on many of the questions on the second reading test.
___	___	7.	I like to read to learn things.
___	___	8.	I like to read when I have extra time.
___	___	9.	Other comments:_____

_____

_____

* Teacher may read questions to the class.